

日本国特許庁
PATENT OFFICE
JAPANESE GOVERNMENT

別紙添付の書類に記載されている事項は下記の出願書類に記載されて
る事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed
this Office.

願年月日
Date of Application:

1999年11月12日

願番号
Application Number:

平成11年特許願第323191号

願人
Applicant(s):

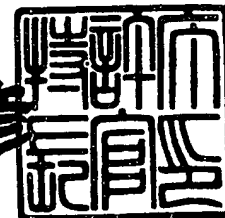
ヤマハ株式会社

RECEIVED
SEP - 7 2000
TECH CENTER 2700

2000年 5月12日

特許庁長官
Commissioner,
Patent Office

近藤隆彦



出証番号 出証特2000-3033130

【書類名】 特許願

【整理番号】 99P119

【提出日】 平成11年11月12日

【あて先】 特許庁長官 殿

【国際特許分類】 H04R 3/00

【発明の名称】 音声処理方法及び装置

【請求項の数】 13

【発明者】

【住所又は居所】 静岡県浜松市中沢町 1 0 番 1 号 ヤマハ株式会社内

【氏名】 藤井 茂樹

【特許出願人】

【識別番号】 000004075

【氏名又は名称】 ヤマハ株式会社

【代理人】

【識別番号】 100092820

【弁理士】

【氏名又は名称】 伊丹 勝

【電話番号】 03-5216-2501

【先の出願に基づく優先権主張】

【出願番号】 平成11年特許願第175925号

【出願日】 平成11年 6月22日

【手数料の表示】

【予納台帳番号】 026893

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9003728

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 音声処理方法及び装置

【特許請求の範囲】

【請求項 1】 少なくとも 1 系統の入力オーディオ信号を複数の信号成分に分離し、前記分離された複数の信号成分の少なくとも一部の信号成分に対して各信号成分毎の個別的な音声処理を行った後、前記複数の信号成分を合成若しくは別々に出力オーディオ信号として出力してなることを特徴とする音声処理方法。

【請求項 2】 前記入力オーディオ信号は、実況放送における環境音成分と実況音声成分とを含み、前記一部の信号成分は、前記環境音成分と実況音声成分であることを特徴とする請求項 1 記載の音声処理方法。

【請求項 3】 前記音声処理は、音場制御処理であることを特徴とする請求項 1 記載の音声処理方法。

【請求項 4】 少なくとも 1 系統の入力オーディオ信号を複数の信号成分に分離する信号分離手段と、

前記信号分離手段で分離された複数の信号成分の少なくとも一部の信号成分に対して各信号成分毎の個別的な音声処理を行う音声処理手段と
を備えたことを特徴とする音声処理装置。

【請求項 5】 前記信号分離手段で分離され、その少なくとも一部が前記音声処理手段で制御された複数の信号成分を、合成若しくは別々に再生系に応じた出力オーディオ信号として出力する出力制御手段

を更に備えたことを特徴とする請求項 4 記載の音声処理装置。

【請求項 6】 前記信号分離手段は、入力オーディオ信号をスペクトル分析して特定の信号成分を抽出し、その抽出された特定の信号成分を元の信号から差し引くことで、残りの信号成分を得るものであることを特徴とする請求項 4 記載の音声処理装置。

【請求項 7】 前記信号分離手段は、入力オーディオ信号に含まれる複数の信号成分のうちの一部を強調し、残りを抑圧する複数の信号強調・抑圧手段を含むものであることを特徴とする請求項 4 記載の音声処理装置。

【請求項 8】 前記信号分離手段は、複数チャネルの入力オーディオ信号の

各チャンネルに対応して設けられ、各チャンネルに対応する信号分離手段は、他のチャンネルのオーディオ信号を補助的に参照して分離精度を高めるものであることを特徴とする請求項 4 記載の音声処理装置。

【請求項 9】 前記音声処理手段は、前記分離された複数の信号成分の少なくとも一部の信号成分の音場制御を行う音場制御手段であることを特徴とする請求項 4 記載の音声処理装置。

【請求項 10】 前記音声処理手段は、前記分離された複数の信号成分の少なくとも一部の信号成分を選択的に消去すると共に、その代替として外部からの入力オーディオ信号を使用するものであることを特徴とする請求項 4 記載の音声処理装置。

【請求項 11】 前記音声処理手段は、前記分離された複数の信号成分の少なくとも一部の信号成分の音質または声質を変更するものであることを特徴とする請求項 4 記載の音声処理装置。

【請求項 12】 前記音声処理手段は、前記分離された複数の信号成分の少なくとも一部の信号成分のピッチを変更するものであることを特徴とする請求項 4 記載の音声処理装置。

【請求項 13】 前記音声処理手段は、前記分離された複数の信号成分の少なくとも一部の信号成分の時間軸に対する速度または話速を変更するものであることを特徴とする請求項 4 記載の音声処理装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

この発明は、入力オーディオ信号に所定の音場制御、音質制御、音程制御等の音声処理を施すことにより、任意のオーディオ信号を得る音声処理方法及び装置に関し、特に複数の発音源を含む入力オーディオ信号の音声処理に適した音声処理方法及び装置に関する。

【0002】

【従来の技術】

従来型の音声処理装置では、1 系統の入力オーディオ信号を 1 系統の音源信号

と仮定し、この入力信号に対して予め定めた処理手順に基づいて、所望の音声処理を行う。音場付加装置を例にとると、従来型の音場付加装置は図 7 のように構成されている。図 7 の装置では、入力された 2 チャンネルステレオ信号のオーディオ信号 (XL, XR) に、音場制御部 101a, 101b で音場制御関数 $\{f(x)\}$ による音場制御処理を施し、音場制御処理された信号 $\{f(XL), f(XR)\}$ を出力制御部 102 で出力制御して出力オーディオ信号 (YL, YR) を出力している。

【 0 0 0 3 】

また、2 チャンネルステレオ信号を入力信号として和信号・差信号を生成し、これらの信号に対して音場制御をするという例では、特公平 7-44759 号公報に開示されているものが知られている。

【 0 0 0 4 】

【発明が解決しようとする課題】

しかし、このような音声処理に際して、事前に入力オーディオ信号を何らかの処理によって複数の信号成分に分離した後、前処理を施し、その前処理で得られた個々の信号成分に対してそれぞれ独立した音声処理を行うような装置は未だ嘗て無く、入力信号に含まれる音源信号を選択的に強調・抑圧したりすることによる自然で臨場感あふれる音場を作り出すことは非常に困難である。

【 0 0 0 5 】

例えば、基本的に発音源が唯一舞台上に存在するということをホール音場の付加装置では仮定している。その付加装置において、付加される初期反射音や残響成分の設定はこの仮定に基づいてなされている。従って、入力オーディオ信号を単一発音源とみなすことができる場合においてのみ、その入力オーディオ信号に対して特に分離、強調・抑圧等の前処理を施さなくても、従来装置で最適な音場付加処理をすることができる。しかしながら、発音源が舞台上以外にも多数存在する場合には、前記仮定に基づいた設定の音場制御では満足な結果は得られない。

【 0 0 0 6 】

つまり、入力オーディオ信号に複数の音場（場所）で収録された音源が含まれる場合においては、従来型の音声処理装置ではそれぞれの音源信号に対して同一

の音声処理がなされるため、出力される結果は必ずしも自然なものになるとは限らないのである。

【 0 0 0 7 】

例えば、スポーツ番組の実況放送音声のように実況音声と観客等の環境音が入り混じっている入力信号に対して従来型装置で反射音・残響音付加を行う場合には、本来環境音にのみホール音場処理を付加したいところであるが、実況音声に対しても反射音・残響音が付加されるため、いわゆる場内放送のような再生音となり非常に不自然なものになってしまうという問題がある。また、従来型装置で音程変更を行うと、実況音声と共に環境音の音程も変化し、聞きづらい再生音となる。

【 0 0 0 8 】

本発明は、このような問題点に鑑みなされたもので、所望とする音声処理の条件がそれぞれ異なる複数の信号成分を含む入力オーディオ信号に対して自然な音声処理が可能な音声処理方法及び装置を提供することを目的とする。

【 0 0 0 9 】

【課題を解決するための手段】

本発明に係る音声処理方法は、少なくとも1系統の入力オーディオ信号を複数の信号成分に分離し、前記分離された複数の信号成分の少なくとも一部の信号成分に対して各信号成分毎の個別的な音声処理を行った後、前記複数の信号成分を合成若しくは別々に出力オーディオ信号として出力してなることを特徴とする。

【 0 0 1 0 】

この発明に係る音声処理方法での前記入力オーディオ信号は、例えば実況放送における環境音成分と実況音声成分とを含み、前記一部の信号成分は、前記環境音成分と実況音声成分である。

【 0 0 1 1 】

また、この発明に係る音声処理方法の前記音声処理は、例えば音場付加などの音場制御処理である。

【 0 0 1 2 】

本発明に係る音声処理装置は、少なくとも1系統の入力オーディオ信号を複数

の信号成分に分離する信号分離手段と、前記信号分離手段で分離された複数の信号成分の少なくとも一部の信号成分に対して各信号成分毎の個別的な音声処理を行う音声処理手段とを備えたことを特徴とする。

【 0 0 1 3 】

また、本発明に係る音声処理装置は、前記信号分離手段で分離され、その少なくとも一部が前記音声処理手段で制御された複数の信号成分を、合成若しくは別々に再生系に応じた出力オーディオ信号として出力する出力制御手段を更に備えるようにしても良い。

【 0 0 1 4 】

この発明に係る音声処理装置に使用される信号分離手段としては、次のようなものが考えられる。即ち、第 1 の信号分離手段は、入力オーディオ信号をスペクトル分析して特定の信号成分を抽出し、その抽出された特定の信号成分を元の信号から差し引くことで、残りの信号成分を得るものである。

【 0 0 1 5 】

また、第 2 の信号分離手段は、入力オーディオ信号に含まれる複数の信号成分のうちの一部を強調し、残りを抑圧する複数の信号強調・抑圧手段を含むものである。

【 0 0 1 6 】

なお、複数チャネルの入力オーディオ信号を処理対象とする場合には、信号分離手段は、例えば各チャネルに対応して設けられ、各チャネルに対応する信号分離手段は、他のチャネルのオーディオ信号を補助的に参照して分離精度を高めるものであることが望ましい。

【 0 0 1 7 】

更に、この発明に係る音声処理装置に使用される音声処理手段としては、次のようなものが考えられる。即ち、第 1 の音声処理手段は、前記信号分離手段で分離された複数の信号成分の少なくとも一部の信号成分の音場制御を行うものである。

【 0 0 1 8 】

第 2 の音声処理手段は、前記分離された複数の信号成分の少なくとも一部の信

号成分を選択的に消去すると共に、その代替として外部からの入力オーディオ信号を使用するものである。

【0019】

第3の音声処理手段は、前記信号成分の少なくとも一部の信号成分の音質または声質を変更するものである。

【0020】

第4の音声処理手段は、前記信号成分の少なくとも一部の信号成分のピッチを変更するものである。

【0021】

第5の音声処理手段は、前記信号成分の少なくとも一部の信号成分の時間軸に対する速度または話速を変更するものである。

【0022】

この発明によれば、スポーツ実況放送等の実況音声成分と環境音成分とが混在したような入力オーディオ信号における音声処理に際して、入力信号を予め複数の信号成分に分離した後、これら信号成分の少なくとも一部にその信号成分に応じた音声処理を施して出力制御することにより、個々の信号成分に最も適した音声処理が可能になるので、違和感のない所望の再生音を創生することができる。例えば、この発明をスポーツの実況放送に適用した場合、環境音と実況音声とを分離して、別々の処理を行うことにより、自然な感じの実況放送を聞き手に対して提供することができるようになる。

【0023】

【発明の実施の形態】

以下、図面を参照して、この発明の実施例を説明する。

図1に本発明の実施例に係る音声処理装置の基本的な構成を示す。

入力オーディオ信号Xは、信号分離部1に入力され、ここで予め決められた方法（アルゴリズム）に従って音源の種類に対応した複数の信号成分に分離される。分離された複数の信号成分（ X_1, X_2, \dots, X_n ）はそれぞれ音声処理部 $2_1, 2_2, \dots, 2_n$ に送られる。なお、この例では音声処理部 $2_1 \sim 2_n$ が信号分離部1で分離された信号成分 $X_1 \sim X_n$ の数だけ設けられている。しかし、所望の処理によっては、入力オ

オーディオ信号 X を無処理 (Through) で出力制御部 3 に送ることも可能である。各音声処理部 $2_1 \sim 2_n$ は、分離された各信号成分に対してその信号成分に適した音声処理を音声処理関数 $\{f_1(X), f_2(X), \dots, f_n(x)\}$ で施し、出力制御部 3 へ音声処理された信号成分 $\{f_1(X_1), f_2(X_2), \dots, f_n(X_n)\}$ を出力する。出力制御部 3 は、音声処理された信号成分を入力信号として、スピーカーの数や配置などの最終出力系に応じてミキシング処理等を施し最終的に出力オーディオ信号 (Y_1, Y_2, \dots, Y_N) を出力する。

【0024】

図 2 に本発明をスポーツの実況放送の音場処理に応用した更に詳細な実施例に係る音声処理方法及び装置のブロック図を示す。

スポーツ実況放送では、入力オーディオ信号にアナウンサーや解説者の実況音声と周囲の環境音とを含む。この入力オーディオ信号 X は、信号分離手段 1 により実況音声 X_1 と環境音 X_2 の 2 つの信号成分に分離される。環境音の信号成分 X_2 は、臨場感を出すために音場制御手段 4 で音場制御され、音場制御された信号成分 $\{f(X_2)\}$ は、出力制御部 3 に入力される。また、実況音声の信号成分 X_1 については、定位感を損わないように無処理とする。出力制御手段 3 では、これらの信号成分 $X_1, f(X_2)$ に、これらを適切に出力するための処理を施して出力オーディオ信号 Y として出力する。

【0025】

図 3 に信号分離部 1 の一実施例を示すが、信号分離に関する手法は分離対象となる入力オーディオ信号に応じて最適な手法を使用すべきであり、本発明では分離手法そのものを限定することはない。入力オーディオ信号 X には、前述したスポーツ実況放送のように実況音声と環境音が混在するような状況を想定している。この信号分離部 1 では、入力されたオーディオ信号からスペクトル分析部 11 によって最初の実況音声成分 X_1 の抽出を行い、次にその抽出された音声成分 X_1 を元の信号 X から差し引くことで環境音成分 X_2 を得る。

【0026】

以下、信号分離処理の流れを図 3 に基づき説明する。

信号分離部 1 に入力されたオーディオ信号 X は、ハイパスフィルタ (HPF) 12

で環境音に含まれる高域成分のみ抽出され、ローパスフィルタ (LPF) 1 3 で実況音声信号成分が含まれる低域成分のみ抽出する。ローパスフィルタ 1 3 から出力される低域信号成分は、ダウンサンプリング部 1 4 でダウンサンプリング処理される。ダウンサンプリング処理は、帯域分割周波数によってサンプリングの比率が異なる。その比率は周波数成分の情報が欠落しない範囲で決定される。例えば、均等 2 分割法では 2 分の 1 の周波数にダウンサンプリングしても、信号成分に含まれる情報は欠落しない。このようにダウンサンプリング処理をすることによって、スペクトル分析部 1 1 における周波数スペクトル分析等の処理量を軽減し、処理速度の向上を図ることができる。

【 0 0 2 7 】

ダウンサンプリング処理された信号成分は、続いて波形切り出し部 1 5 で適当な時間窓を掛けられて波形の切り出しをされる。その後、切り出された信号成分はスペクトル分析部 1 1 へ送られる。スペクトル分析部 1 1 に入力された信号成分は、初めに FFT 部 1 6 で周波数領域の信号成分に変換される。この際、本実施例でのスペクトル分析部 1 1 では変換手法として高速フーリエ変換 (FFT) を用いた構成をとっている。しかしながら、本発明においては特にこの方法に限定しなくてもよい。ただし、この実施例における時間一周波数変換された信号成分は、各周波数成分の周波数情報とその周波数成分の強度情報で表現されていることと定義する。

【 0 0 2 8 】

次に、FFT 部 1 6 で変換された周波数領域の信号成分に対して、実況音声成分の抽出並びに同定判定を調和成分抽出部 1 7 及び音源同定判定部 1 8 で行う。基本的に音声信号は、基本波に対して整数倍の高調波成分が現れる調和構造を有しているために、この抽出及び同定判定処理では前記調和構造を有する信号成分が前記周波数領域の信号成分に存在するか否かを判定する。判定方法に関しては、スペクトラム・サマメーション (Spectrum Summation) 法などを利用することが可能である。抽出及び同定判定処理の結果で、実況音声成分が周波数領域の信号成分に存在すると判定された場合には、実況音声成分の基本波を含む高調波成分の各周波数及び強度の情報が同定される。

【 0 0 2 9 】

ただし、抽出及び同定判定処理で同定された信号成分は、この段階では本来の実況音声成分とその実況音声成分の高調波成分に対して同じ周波数の環境音成分も重畳されたものであるから、前記環境音成分を除去する必要がある。理論的にはこのような同一周波数成分を完璧に分離することは不可能であるが、本実施例では環境音スペクトル包絡推定部 2 0 において環境音成分のスペクトル包絡（周波数特性）は時間的にほぼ一定との仮定に基づき、周波数特性のパワー変化を入力オーディオ信号の瞬時パワー並びに帯域分割された高域信号成分の瞬時のパワーから推定している。また、環境音スペクトル包絡推定部 2 0 では環境音成分の平均的なスペクトル包絡を、予め保持するスペクトル包絡情報と音声信号なしと判断されたときの環境音信号のスペクトル包絡とを統計的に計算処理することで得ている。

【 0 0 3 0 】

調和成分抽出部 1 7 と音源同定判定部 1 8 を通して出力される周波数成分から環境音スペクトル包絡推定部 2 0 で推定された周波数成分を、スペクトラム・サブトラクション (Spectrum Subtraction) 法などを用いてスペクトル減算部 1 9 において差し引くことにより、実況音声信号成分（分離対象信号）は得られる。差し引かれた周波数成分（分離対象信号）は、逆FFT部 2 1 に送られ周波数領域の信号成分から時間領域の信号成分に変換される。変換された信号成分はアップサンプリング部 2 4 へ送られアップサンプリング処理を施されて、最終的に元のサンプリング周波数に戻される。戻された信号成分は、実況音声信号成分X1として音場制御部 4 へ出力される。

【 0 0 3 1 】

また、一方でスペクトル減算部 2 2 において、FFT部 1 6 から出力された信号成分から前記実況音声成分（分離対象信号）を周波数領域で減算する。その結果求められた信号成分を逆FFT部 2 3 により逆FFT処理することによって時間領域の信号成分に戻す。時間領域に戻された信号成分にハイパスフィルタ 1 2 を経てきた高域信号成分を加算器 2 5 で加算することにより、環境音信号成分X2を得ることができる。最終的にこうして得られた環境音信号成分X2を、実況音声信号成分

X1とは別系統から音場制御部4に出力する。信号分離部1ではこのような構成により入力オーディオ信号を複数の信号成分に分離している。

【0032】

この実施例においては、一般的に実況音声は明瞭で聞き取り易くなければならないとすることを考慮して、実況音声成分には反射音や残響成分等の音場付加は行わないか又は極僅かに付加するだけとしている。環境音成分に対しては、音場全体を包み込むような臨場感を付加するために、適度に反射音や残響成分を良く知られている仮想音像定位処理等の手法により付加する。このような音場制御部4での音場制御は、言うまでもなく入力オーディオ信号の性質やユーザーの要求に大きく依存するものであり、制御方法そのものは限定される限りではない。

【0033】

なお、上記実施例では、信号分離部1として入力オーディオ信号をスペクトル分析して特定の信号成分を抽出するようにした。これに対し、図4に示すように、信号分離部1として信号成分毎に信号強調・抑圧処理を行うように構成することもできる。図4の装置では、信号分離部1が複数の信号強調・抑圧部 $31_1 \sim 31_n$ で構成されている。信号分離部1に入力されたオーディオ信号Xは、予め決められた方法（アルゴリズム）に従って複数の信号成分に分離される。分離された複数の信号成分のうち信号強調・抑圧部 $31_1 \sim 31_n$ で強調・抑圧処理された信号成分(X_1, \dots, X_n)はそれぞれ音声処理部 $2_1 \sim 2_n$ に送られる。この時、基本的に音声処理部 $2_1 \sim 2_n$ は信号強調・抑圧部 $31_1 \sim 31_n$ の出力信号の数だけ設けられる。また、所望の処理によっては、信号強調・抑圧部 $31_1 \sim 31_n$ を通さずに無処理(Through)で出力制御部3に送ることも可能である。この音声処理部 $2_1 \sim 2_n$ では、各強調・抑圧処理された信号成分に対して所定の音声処理を施す。そして、出力制御部3に音声処理された信号成分 $\{f_1(X_1), \dots, f_n(X_n)\}$ を出力する。出力制御部3は、音声処理された信号成分を入力信号として、再生系に応じたミキシング処理等の出力制御を施してから、出力オーディオ信号(Y_1, Y_2, \dots, Y_N)として出力する。

【0034】

なお、以上は1入力の場合について説明したが、図5に示すように入力オーデ

ィオ信号を 2 系統 (XL, XR) とした場合には、左入力オーディオ信号 (XL) に対しては左信号分離部 (L) 1a が、また右入力オーディオ信号 (XR) に対しては右信号分離部 (R) 1b がそれぞれ設けられる。各信号分離部 1a, 1b で各入力オーディオ信号を複数の信号成分 (XL1, ..., XLn, XR1, ..., XRn) に分離する。その後、音声処理部 2a1 ~ 2an, 2b1 ~ 2bn で、分離された信号成分に対して個別に音声処理を施し、更に出力制御部 3 で出力系に応じた出力制御処理をしてから出力する。このような複数入力の場合では、中央定位成分等の主要成分が各入力中に共通に含まれることが多く、例えばこの共通の成分そのものが信号分離すべき対象成分となるようなときは、これら各入力同士をレベル調整し対象成分レベルをほぼ同一レベルとしたうえで加減算することにより、簡単かつ比較的精度良く分離（分離のための打消除去等）できる。すなわち、信号分離の精度向上を目的として、図 5 に破線で示すように補助入力信号 (XLs, XRs) を、互いに他方の入力信号として補助的に利用することができる。なお、この場合も、信号分離処理の主対象は入力オーディオ信号であることに変わりなく、この発明の趣旨に沿うものであることは言うまでもない。

【 0 0 3 5 】

図 6 は、上述した音声処理装置をスポーツ実況放送の音場処理に適用した更に詳細な実施例を示す図である。基本的な入出力オーディオ信号は、左チャンネル入力オーディオ信号 (XL) と右チャンネル入力オーディオ信号 (XR) の 2 チャンネルステレオ入力である。信号成分は、所望の設定で典型的なスポーツ番組の音声を想定して、センターに左実況音声成分 (XLsp) と右実況音声成分 (XRsp) が定位で配置され、左環境音成分 (XLse) と右環境音成分 (XRse) がある程度の広がりを持って配置されているものと仮定する。

【 0 0 3 6 】

先ず、信号分離部 1 は入力信号の数に基づき設けられるので、この場合左信号分離部 (L) 1a と右信号分離部 (R) 1b の 2 系統が設けられる。信号分離部 1 では、前記左右入力オーディオ信号 (XL, XR) に対して各入力信号毎に内部処理が行われ、実況音声成分と環境音成分とに入力オーディオ信号を分離する。例えば、左入力オーディオ信号 (XL) の場合をみると、左信号分離部 (L) 1a で左実況音

声成分 (XLsp) と左環境音成分 (XLse) とに前記信号を分離するという具合である。内部の信号分離処理そのものは、入力されたモノラル信号に対して行われる。この際、本実施例のように 2 チャンネルステレオ入力信号でセンターに同一音源が含まれているような場合には、左右入力オーディオ信号のお互いの入力信号を前述の補助入力信号 (XLs, XRs) として破線で見られるように補助的に利用し、分離対象信号成分の強調処理などを行うことで、入力オーディオ信号の分離精度を高めることも可能である。その後、分離された信号成分の数に応じて設けられた音場制御部 4 で、各実況音声成分と環境音成分は所定の音場制御を施される。

【0037】

音場制御部 4 は、実況音声成分に対する音場制御関数 $\{f(x)\}$ をもつものと環境音成分に対する音場制御関数 $\{g(x)\}$ をもつものとが設けられ、音場制御部 4 では各成分に所定の音場制御がなされる。音場制御部 4 で制御された左右各実況音声成分 $\{f(XLsp), f(XRsp)\}$ と環境音成分 $\{g(XLse), g(XRse)\}$ は、出力制御部 3 へ送られる。出力制御部 3 では初めに、音場制御部 4a1, 4b1 から送られてきた左右実況音声成分を加算器 4 1 で合成する。次に、加算器 4 1 及び乗算器 4 4 で合成された実況音声成分に右環境音成分 $\{g(XRse)\}$ を加算器 4 3 によって合成し、加算器 4 1 及び乗算器 4 5 で合成された実況音声成分に、左環境音成分 $\{g(XLse)\}$ を加算器 4 2 によって合成する。これにより、再生系に合わせたかたち、例えば、左右出力オーディオ信号 (YL, YR) として 2 チャンネルステレオ再生で出力する。

【0038】

なお、図 6 では、その再生系として 2 チャンネルステレオ出力 (YL, YR) を想定しているが、本発明では再生系そのものを限定することはない。一般的に出力チャンネルを多チャンネル化すればするほど音場の臨場感は増すと言われているが、更に出力系統を増やす場合には、音場制御部 4 も増加分に対応する形で追加したり変更したりして出力を増やすなどの応用処理が必要となることは言うまでもない。ここでは、前述のように実況音声を中心にセンターに定位させ、環境音を左右に配置する出力を想定している。

【0 0 3 9】

また、この発明における各信号成分毎の音声処理は、上述した音場制御処理に限定されるものではない。例えば、アナウンサー 1 人と解説者 a, b 2 人の計 3 人でのスポーツ実況放送に応用した場合、アナウンサーの実況音声に対しては所望の音程または声質に変更し、解説者 a の実況音声に対しては消去し、解説者 b の実況音声に対しては話速を変更したりする等の音声処理を施すことができる。

【0 0 4 0】

また、聴覚機能の劣る高齢者や障害者用としては、単に音量を上げるだけでなく、その明瞭性向上処理（特に高域成分の強調）も有用である。一方、環境音に対しては、音量調整や音質（イコライジング）等が有用である。こうした制御は、入力オーディオ信号の性質、また、ユーザーの好みに依存するものであり、その制御方法そのものはこれに限定するものではない。最終的に、音声処理部 2 で処理された信号は、出力制御部 3 へ送られる。

【0 0 4 1】

更に、この発明における音声処理には、分離された信号成分を選択的に消去し、代わりに外部からの入力信号を使用する処理も含まれる。

【0 0 4 2】

なお、上述の実施例では、その再生系として 2 チャンネルステレオ出力（YL, YR）を想定しているが、本発明では再生系そのものを限定することはしない。一般的に出力チャンネルを多チャンネル化すればする程音場の臨場感は向上すると言われているが、更に出力系統を増やす場合には、音声処理部 2 も増加分に対応する形で追加したり変更したりして出力を増やす等の応用処理が必要となることは言うまでもない。ここでは、前述のように実況音声を中心に定位させ、環境音を左右に配置する出力を想定している。

【0 0 4 3】

【発明の効果】

以上述べたようにこの発明によれば、音声処理に際して事前に入力されたオーディオ信号を、複数の信号成分に分離してから個別に音声処理するため、所望とする再生音を創生することができる。

【図面の簡単な説明】

【図 1】 本発明の実施例に係る音声処理装置の構成を示す図である。

【図 2】 同音声処理装置の更に詳細な実施例を示すブロック図である。

【図 3】 本発明の実施例に係る信号分離部の一実施例の処理を説明するための図である。

【図 4】 信号分離部に信号強調・抑圧処理を用いた他の実施例の基本構成を示す図である。

【図 5】 同信号分離処理の本発明を 2 チャンネル信号に適用した他の実施例を示す図である。

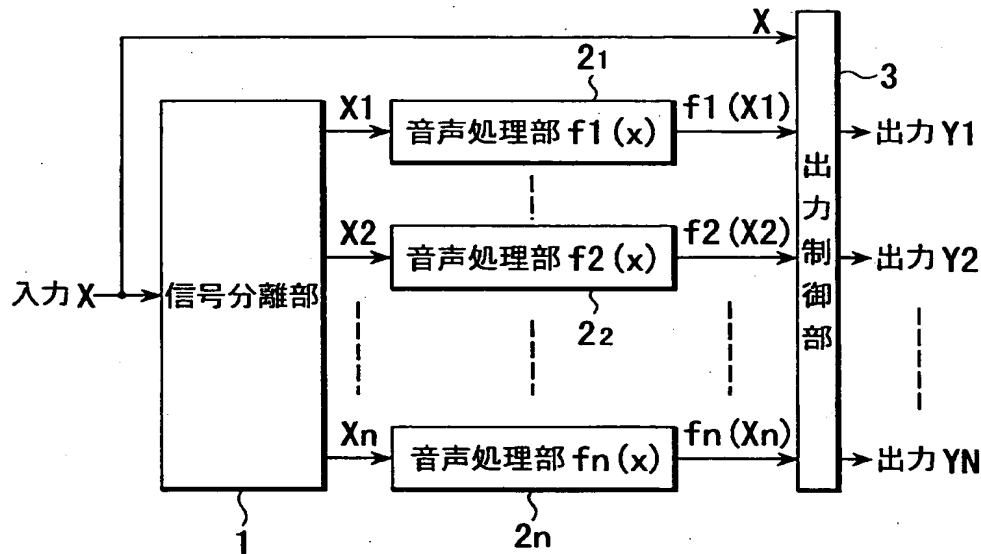
【図 6】 同実施例の更に具体的な音声処理装置の構成を示すブロック図である。

【図 7】 従来例の音声処理装置の構成を示す図である。

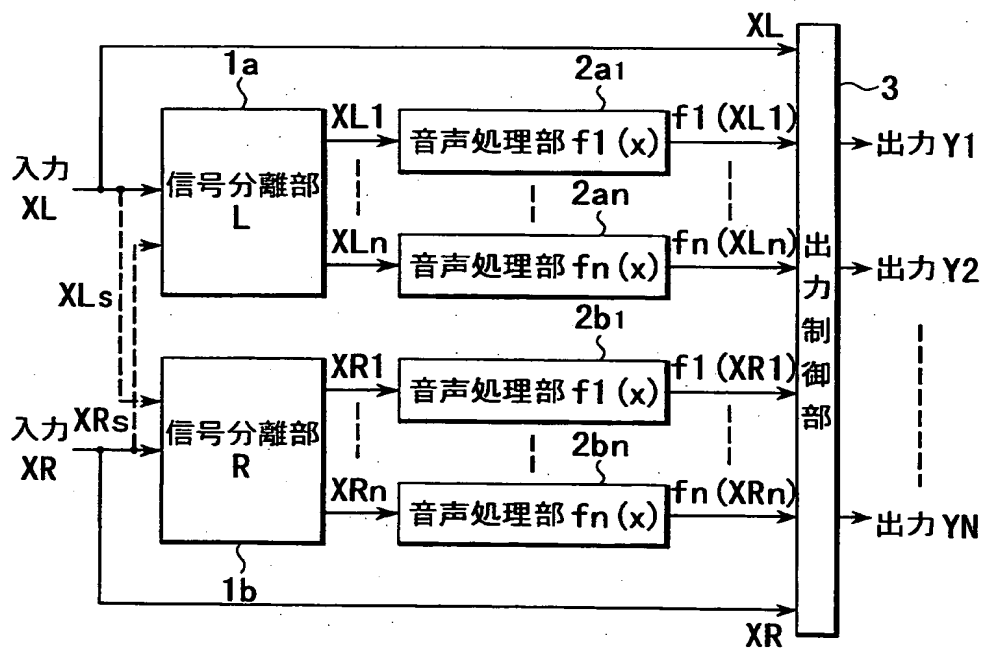
【符号の説明】 1 … 信号分離部、 2 … 音声処理部、 3 … 出力制御部、 4 … 音場制御部、 1.1 … スペクトル分析部、 3.1 … 信号強調・抑圧部。

【書類名】 図面

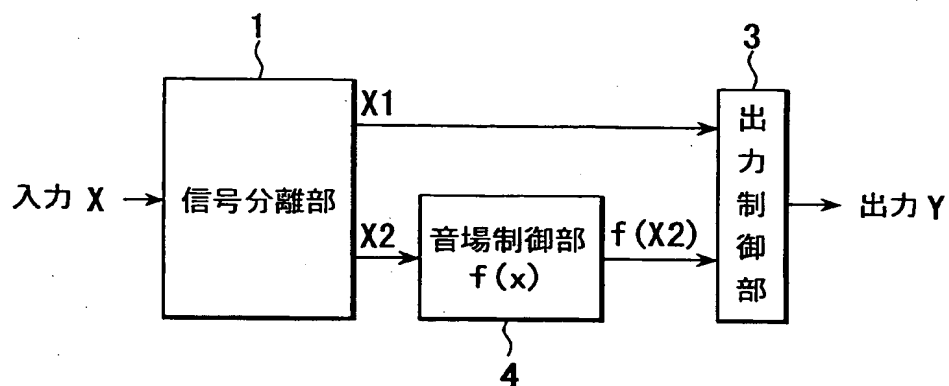
【図 1】



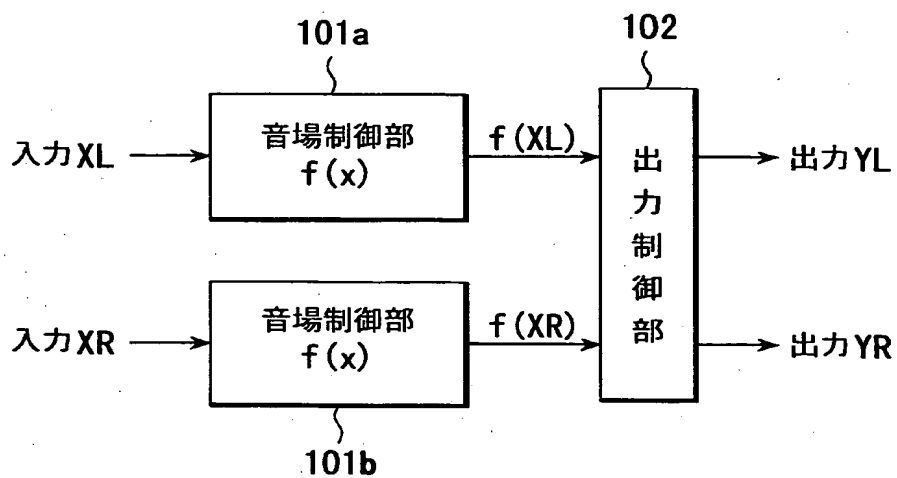
【図 5】



【図 2】

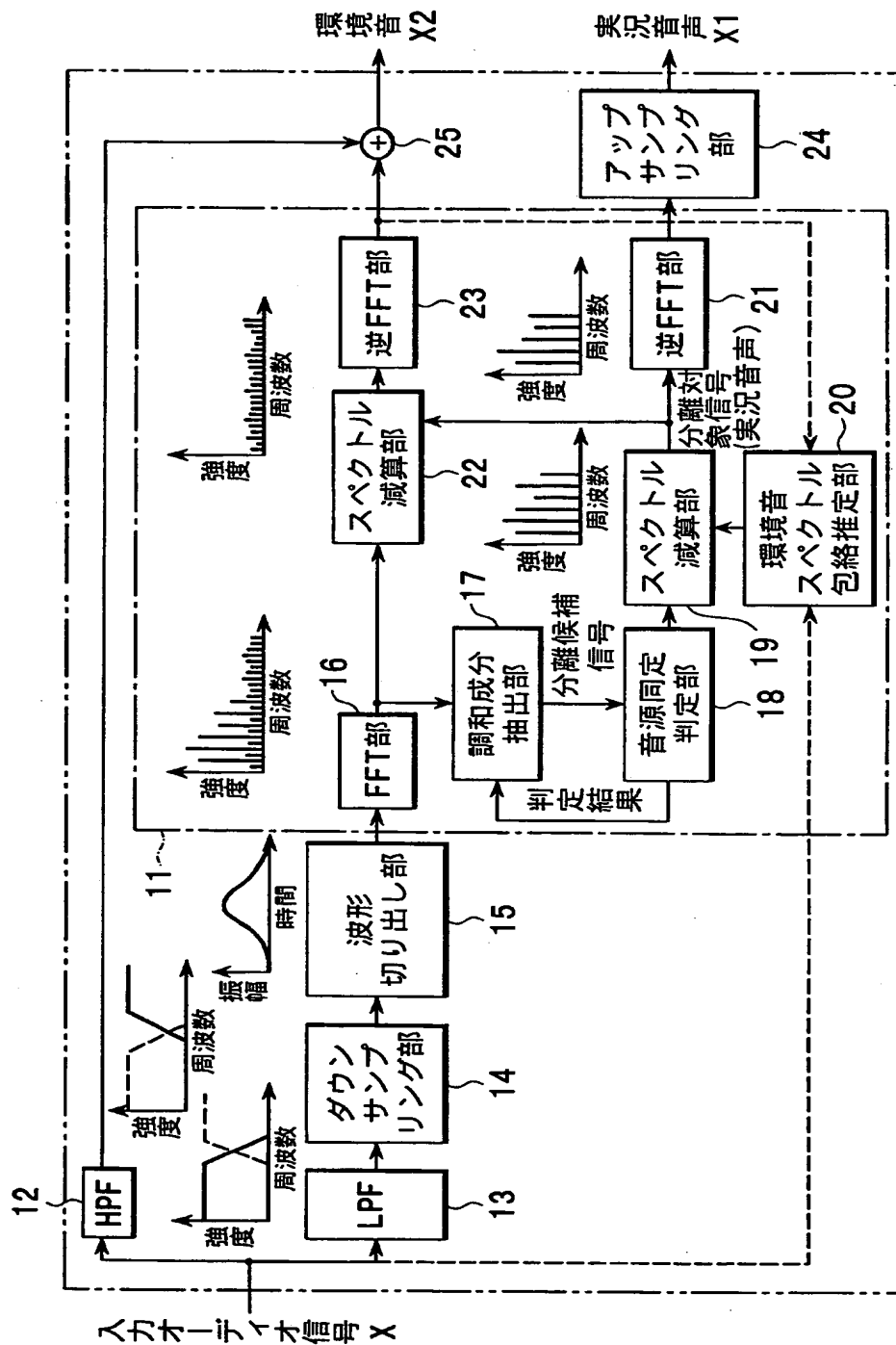


【図 7】

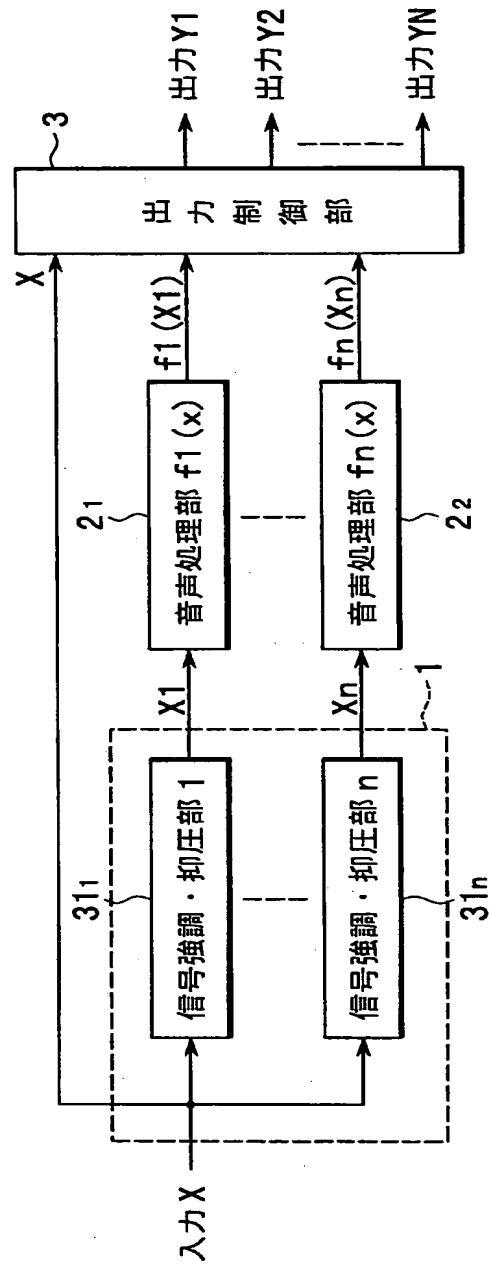


XL : 左入力信号
 XR : 右入力信号
 $f(x)$: 音場制御関数
 YL : 左出力信号
 YR : 右出力信号

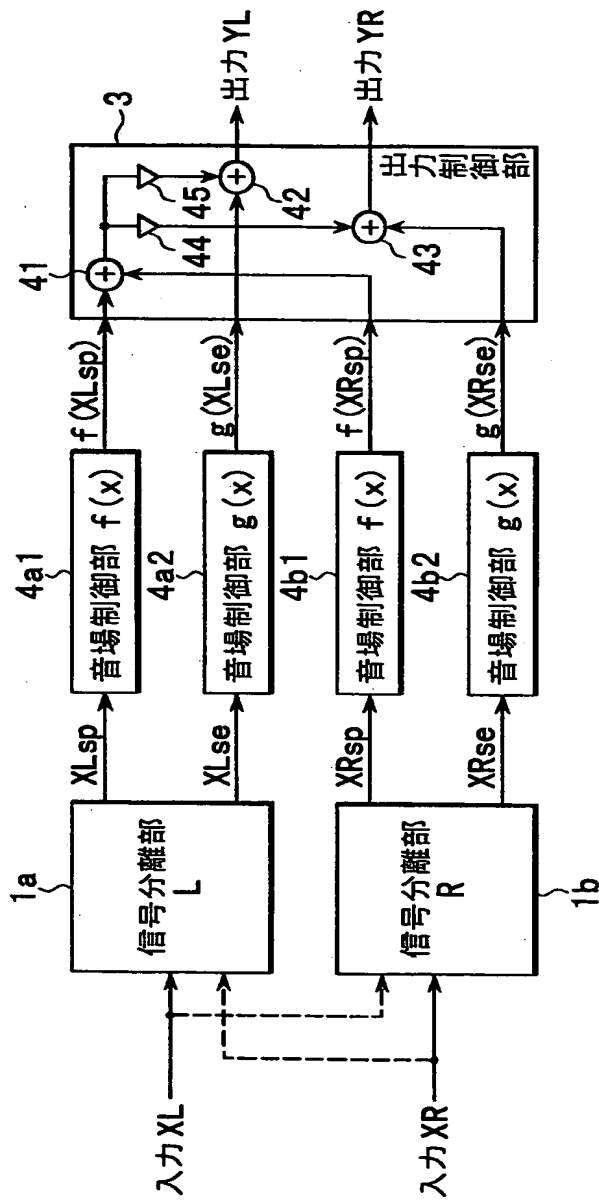
【図 3】



【図 4】



【図 6】



XL, XR: 入力信号
 XLsp, XRsp: 分離された実況音声
 XLse, XRse: 分離された環境音
 $f(XLsp)$, $f(XRsp)$: 音場制御された実況音声
 $g(XLse)$, $g(XRse)$: 音場制御された環境音
 YL, YR: 出力信号

【書類名】 要約書

【要約】

【課題】 音場付加処理に際して事前に入力されたオーディオ信号を、複数の信号成分に分離してから音声処理し、自然な臨場感に富んだ音声の創生ができるようにする。

【解決手段】 少なくとも1系統の入力オーディオ信号を予め複数の信号成分に分離し、分離された複数の信号成分の少なくとも特定の信号成分に対して所定の音声処理を行ったのち、前記複数の信号成分を合成若しくは別々に出力オーディオ信号として出力してなる音声処理装置により、所望とする音声再生性能を備える。

【選択図】 図5

出 願 人 履 歴 情 報

識別番号 [0 0 0 0 0 4 0 7 5]

1. 変更年月日	1 9 9 0 年 8 月 2 2 日
[変更理由]	新規登録
住 所	静岡県浜松市中沢町 1 0 番 1 号
氏 名	ヤマハ株式会社